REPORT DOCUMENTATION PAGE Dist. A

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
		ANNUAL 01 Jan 93 TO 31 Dec 93

4. TITLE AND SUBTITLE

ERRORS IN SKILLED PERFORMANCE

5. FUNDING NUMBERS

F49620-93-1-0094

61102F

2313/BS

6. AUTHOR(S)

Dr Dan Woltz

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Dept of Educational Psychology
University of Utah
327 Milton Bennion Hall
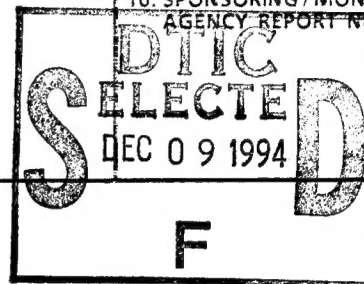Salt Lake City, UT 84112

8. PERFORMING ORGANIZATION REPORT NUMBER

AEOSR-TR· 94 0754

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR/NL
110 Duncan Ave Suite B115
Bolling AFB DC 20332-0001

Dr Tangney

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

DTIC
SELECTED
DEC 0 9 1994
F

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release;
distribution unlimited.

12b. DISTRIBUTION CODE

A

1994 1201 114

13. ABSTRACT (Maximum 200 words)

This report details work from the first year of a two year project on undetected errors in skilled cognitive performance. First, we present the theoretical background in which our research program is cast: Anderson's (1983) ACT* theory of memory and skill acquisition. Second, we outline a set of circumstance that we believe promote error making -- namely, long-term priming (i.e., training on a subset of the universe of possible problem solution types), short-term priming (i.e, presenting multiple surface structure instantiations of a single, deep structure problem type in succession), and working memory load (i.e., presenting a concurrent secondary task that requires working memory capacity). These form the basis of our empirical work (presented later). Third, we describe our methodology for "detecting: undetected errors -- that is, how we are able to decide whether an error is in or out of a subject's's awareness. Fourth, we present our empirical work. One set of three studies finds evidence for "general processing sequence" memory. This is memory for the order of application of cognitive operations, and is independent of the actual items (i.e., surface structures of problems) subjects actually encountered.

14. SUBJECT TERMS			15. NUMBER OF PAGES
			16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
(U)	(U)	(U)	(U)

Errors in Skilled Performance
Annual Technical Report
for the Period 1 January 1993 through 31 December 1993
for Grant #F49620-93-1-0094 (Cognition Program)
Dan J. Woltz and Michael K. Gardner
University of Utah

Submitted to:
US Air Force Office of Scientific Research
Air Force Systems Command, USAF
Bollling AFB, DC 20332-6448

Summary

This report details work from the first year of a two year project on undetected errors in skilled cognitive performance. First, we present the theoretical background in which our research program is cast: Anderson's (1983) ACT* theory of memory and skill acquisition. Second, we outline a set of circumstance that we believe promote error making -- namely, *long-term priming* (i.e., training on a subset of the universe of possible problem solution types), *short-term priming* (i.e., presenting multiple surface structure instantiations of a single, deep structure problem type in succession), and *working memory load* (i.e., presenting a concurrent secondary task that requires working memory capacity). These form the basis of our empirical work (presented later). Third, we describe our methodology for "detecting" undetected errors -- that is, how we are able to decide whether an error is in or out of a subject's awareness. Fourth, we present our empirical work. One set of three studies finds evidence for "general processing sequence" memory. This is memory for the order of application of cognitive operations, and is independent of the actual items (i.e., surface structures of problems) subjects actually encountered. A second set of two studies demonstrates that this general sequence memory is procedural in nature, and is not open to conscious introspection. A third set of two studies supports the notion that general sequence memory, in conjunction with the long-term priming paradigm, can produce undetected errors. A fourth set of three studies explored undetected errors in the short-term priming paradigm. No evidence was found to support short-term priming as a mechanism in the production of undetected errors. Finally, a study is presented using a working memory load. It found that working memory load increases the likelihood of undetected errors in moderately skilled individuals, but reduces the likelihood of undetected errors in highly skilled individuals. Future directions for the research include the following: (1) studying two new task domains: a procedural learning number classification task and a water quality index computation task (these are currently being programmed); (2) studying the effects of forcing subjects to temporarily modify well-learned procedures; (3) looking at individual differences in the susceptibility to undetected error making, and how to predict it; and (4) computer simulation of the partial matching mechanism we posit underlies undetected error making in highly skilled subjects.

Research Objectives

A general finding in the literature is that higher levels of skill in cognitive tasks result in faster and more error free performance (e.g., Bryan & Harter, 1899; Crossman, 1959; Fitts & Posner, 1967; LaBerge, 1973; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). While this trend is true in general, recent theoretical advances in cognitive science (e.g., Anderson, 1983, 1987) lead to the prediction that experts (i.e., highly skilled performers) will be more error prone under certain training/transfer circumstances than novices. Furthermore, these errors should be unavailable to conscious introspection -- that is, they should be undetected by the performer. A series of experiments were performed to test these predictions.

Theoretical Background

The predictions concerning the relative quantity of errors by skilled performers, and their ability to detect these errors, derive from the ACT* theory of skill acquisition proposed by John Anderson (1983, Singley & Anderson, 1989). In Anderson's theory, cognitive skills are represented as a set of *productions*, or condition-action statements, along with a hierarchical *goal structure* that facilitates problem solving. Skill acquisition starts with the so-called weak problem solving methods (e.g., analogy, means-ends analysis, etc.) that are applicable across a wide range of problem types. The weak methods take declarative knowledge stated in the problem instructions, and organize the initial productions necessary to solve the problem. At first these productions are both numerous and general. Two processes restructure these original productions into a smaller number of more task specific ones. The first process is called *proceduralization*. This process removes variables from the production definitions, and embeds task specific constants in their place. Proceduralized productions are hypothesized to be more efficient, but less flexible, than their generalized counterparts. The second process is called *composition*. Composition takes two or more productions which fire in sequence and collapses them into a single, more complex production. Again, this process leads to greater efficiency by reducing the number of productions necessary to represent task performance, but at the cost of flexibility. Their may be other costs as well, such as the number of conditions that must be active in working memory at one time in order to fire the production.

Productions are assumed to be learned in an all-or-none fashion. However, productions have a "strength" associated with them. Each time a production fires, its strength is enhanced, making it more likely to fire again in the future.

Circumstances That Promote Errors

Certain task circumstances promote error making. In one paradigm, which we have designated *long-term priming*, subjects learn a set of rule sequences for performing a task requiring the serial application of separate cognitive processing steps. The set of rule sequences learned during training, however, constitutes only a subset of the total universe of rule sequences. During transfer subjects are exposed to the entire universe of rule sequences (or, occasionally, just a larger subset of the universe). We believed that subjects will be more likely to make errors on those rule sequences that are new to the transfer session, and that they will be relatively unaware of these errors. Furthermore, this tendency will be greater among more skilled performers (i.e., performers with greater degrees of expertise in the task). This prediction derives from two factors: (1.) expert subjects will have more highly composed productions than novice subjects, thus reducing their ability to introspect on how they solve the task (therefore, the errors are undetected); and (2.) expert subjects will have composed productions for rule sequences learned during training with relatively high strength, and, therefore relatively low firing thresholds. This should lead to occasions when these *strong-but-wrong* productions fire inappropriately to new rule sequences that achieve a partial match to the old productions' conditions. Although the match is only partial, it is compensated for by the extremely low firing threshold of these highly-practiced productions.

It is also possible that seeing the exact same instances of a rule-sequence would result in a benefit to processing latency and accuracy. Such *instance effects* would be predicted as part of the process of proceduralization (described above). To calculate instance effects, one must compare latency and/or error rate during transfer for old items (that is, items with a previously practiced rule sequence and previously seen surface structure) to new instances of old items (that is, items with a previously practiced rule sequence, but with a different surface structure).

Another paradigm, which we have designated *short-term priming*, is similar to long-term priming in its rationale; however, the strengthening of productions is hypothesized to occur over a brief time frame: say several trials within a trail block. If either the repeated firing of a production temporarily strengthens it, or if the presentation of contextual information associated with the firing of a production spreads activation

that temporarily lowers the firing threshold of the production, then temporary priming should be possible. Again, we would expect errors in situations where a temporarily primed production shared a partial match to the conditions of an item. Furthermore, this effect should be more pronounced in experts, and should be more likely to be undetected by experts.

A final task circumstance that should promote errors is the addition of a *working memory load*. A working memory load should promote errors for several reasons. First, for novices, the memory load should compete for working memory capacity with the productions necessary to complete the experimental task. To the extent that their is insufficient working memory capacity to complete both tasks, one or both tasks should suffer. Second, for experts, composed productions require more information to be held in working memory to satisfy their conditions. For example, if production A requires condition 1, and production B requires condition 2, the composition of these productions, production A-B, requires conditions 1 and 2 to be present for its firing. This additional working memory load is offset, to some degree, by requiring fewer productions to complete a task, and by simplifying the task's goal structure. Thus it is difficult to make specific predictions concerning experts' performance.

Detecting Undetected Errors

A methodological problem faced in this research program was how to decide if a particular error made by a subject was a "detected" error (an error in the subject's awareness which could be corrected) or an "undetected" error (an error outside of the subject's awareness). To make the distinction, we designed our tasks in the following way: during training, subjects were instructed to perform as quickly as possible, while maintaining an accuracy rate of 90%. If a response on a trial was incorrect, the word WRONG and a low tone was presented for 2 s. At the end of blocks, subjects were given feedback on median latency and percentage of errors made during that block of trials (and all previous blocks during that day's session). Subjects were thus encouraged to decrease their latency until this pushed their error rate above 10%. During transfer, subjects were given new instructions. They were told to attempt to attain an accuracy rate of 100%. In order to achieve this difficult new goal, subjects were told they could "retake" any trial they thought they had made an error on by pressing the spacebar on the computer keyboard. By retaking the trial, only data from the "corrected" trial would count toward their performance goal. In fact, we collected data from all trials, and were consequently able to distinguish undetected from detected errors. Indeed, four possibilities exist: correct trials thought to be correct by the subject, correct trials thought to be an error by the subject, incorrect trials thought to be incorrect by the subject (i.e., detected errors), incorrect trials thought to be correct by the subject (i.e., undetected errors).

The methodology was tested in a pilot study. The study included 40 subjects. The subjects learned a task we have designated *number reduction*. On each trial, a 4-digit number was presented that had to be reduced to a single digit. This reduction was accomplished by applying some combination of the following component rules. The *same rule* stated that two identical numbers could be reduced to a single digit of that same number (e.g., 77=7). The *midpoint rule* stated that two numbers that differed by two could be reduced to their midpoint (e.g., 53=4). The *contiguous rule* stated that two numbers in either an ascending or descending sequence could be reduced to the next number in the sequence (e.g., 32=1; 67=8). Finally, the *last rule* stated that two number whose difference was greater than 2 could be reduced to the last of the two numbers (e.g., 28=8; 63=3). These rules were applied to multi-digit stimuli by parsing the stimuli pairwise left to right and carrying forward intermediate solutions to be combined with the next digit in the stimulus (e.g., 9687=7).

Subjects were also administered a brief questionnaire that asked the following questions: (1.) Did you use the spacebar to correct errors today? (2.) Did the spacebar ever fail to operate properly at any time during the session? and (3.) Were you ever aware of an error, but decided not to correct it?

Results from the pilot study yielded the following conclusions and modifications. Subjects overwhelmingly said they used the spacebar for error correction and that it operated properly. Only occasionally did a subject indicate that they were aware of an error, but decided not to correct it. When

this did occur, it was often because the awareness of an error did not occur until the following trial, at which time it was no longer possible to correct the error. The error detection software was modified to allow correction of an error not only during the interstimulus interval, but also during the following trial (until a numeric response was given). This allowed subjects to correct the vast majority of detected errors.

Long-Term Priming

Several studies investigated the effects of long-term priming on error making, error awareness, and latency. The first two studies used a simplified version of the number reduction task. In this paradigm, subjects are given two rules for reducing multi-digit stimuli containing combinations of the numbers 1, 2, and 3. The *same rule* states that if two adjacent digits are the same, they may be reduced to a single digit of that value (e.g., 11=1). The *different rule* states that if two adjacent digits are different, they can be reduced to the unused number (e.g., 13=2, 23=1). The rules are applied stepwise from left to right, carrying forward intermediate results. Stimuli consisted of three digit triplets which yielded a single digit solution.

Study I Objectives. Study I was designed to investigate the roles of memory for general processing sequences and memory for specific instances on latency and error rate in the simplified number reduction task.

Study I Methods. In study I, subjects consisted of 479 US Air Force recruits in their eleventh day of basic training at Lackland AFB, San Antonio, TX. Of these, 77 were eliminated from the final data analysis, because their data indicated lack of effort. This left a final pool of 402 subjects between the ages of 17 and 27.

The training phase consisted of eight blocks of 24 trials. Each subjects performed 12 unique instances during training. However, subjects differed with regard to whether those 12 instances came from *four* different rule sequences, or whether they came from only *two* rule sequences.

The transfer phase also consisted of eight blocks of 24 trials. However, during transfer, all subjects were exposed to all 24 possible stimuli (each of the four possible rule sequences had six separate exemplars).

Study I Results. Training data showed no difference between the two and four rule conditions in terms of error rate. However, the conditions did differ with respect to latency, with the two rule condition being significantly faster than the four rule condition. This result reflects the importance of memory for general processing sequences, since both groups received equal amounts of practice with each component rule, and equal numbers of unique exemplars.

In transfer, subjects saw both old trials (previously presented instances) and new trials (never before seen instances). For the four rule sequence group, all new trials represented previous used rule sequences. For the two rule sequence group, however, new trials represented both new instances and new rule sequences. Comparing new and old trials for the four sequence group gives an estimate of the instance effect, or the effects of proceduralization. Differences were relatively small (on the order of 50 msec in latency and 1% in error rate), but statistically significant. Comparing new and old trials for the two sequence group gives an estimate of the of the effects of memory for general processing sequences, or composition with variable arguments. This effect was sizable (on the order of 300 msec in latency and 5% in error rate) and also statistically significant. Thus memory for general processing sequences is an important component in learning a task such as number reduction, and is more powerful an influence than memory for individual instances. This is in accord with predictions made by some theories of skilled performance (e.g., Anderson, 1983; and MacKay, 1987), but in contrast to predictions made by some instance-based theories (e.g., Logan, 1988).

Study II Objectives. Study II introduced differing levels of expertise or practice to the design of Study I. This allowed us to study the time course of the development proceduralization and composition. Some evidence (e.g., Anderson, 1983; Frensch, 1991) exists to suggest that such effects develop quite quickly,

although it would seem reasonable to propose that these changes in performance are gradual. A second objective was to investigate negative transfer in the simplified number reduction task. Subjects exposed to new rule sequences during transfer actually performed worse on these than they performed on new sequences at the start of training (20% errors versus 10% errors). This suggests one source of error making in skilled cognitive performance: the carrying over of previously learned general rule sequence memory to new situations in which these sequence are inappropriate.

Study II Methods. Methods were identical to those of Study I, except for a between subjects manipulation of amount of training. Subjects received either 1, 2, 4, or 8 blocks of practice during training. Subjects were 796 US Air Force recruits in their eleventh day of basic training at Lackland AFB, San Antonio, TX. Of these, 129 were eliminated from the final data analysis, because their data indicated lack of effort. This left 667 subjects for the final data analysis.

Study II Results. Latency data over training blocks were extremely similar for the four skill level groups (i.e., 1, 2, 4, or 8 blocks of trials during training). On the one block common to all groups, there was no significant difference among the groups. Transfer performance was examined both in terms of latencies and error rates. Latency data revealed two things: (1.) skill level affected transfer latency (the highest skill group was approximately 200 msec faster than the lowest skill group), and (2.) general sequence memory effects were a function of skill level (greater degrees of practice lead to larger differences between old and new sequence trials). Error data also showed greater general sequence memory effects with higher levels of skill; that is, high skill levels lead to larger differences in error rates between old and new trials, with subjects being more error prone on the new trials. Finally, there was evidence for the growth of negative transfer with skill. Subjects with more practice during training made more errors on new trials (in an absolute sense) than subjects with less practice.

Thus Study II provided evidence for the gradual growth of composition with skill, and for the development of negative transfer in close transfer situations along with skill that facilitates correct performance on old trials. One question still remained: would the effects observed in the simplified number reduction task generalize to more complex tasks, such as the full fledged number reduction task?

Study III Objectives. The objective of study III was to replicate the findings of study I within the more complex task environment of full fledged number reduction (described above).

Study III Methods. In this version of number reduction, there were 24 possible rule sequences with each sequence being three rules long. A within subjects design was used in which each subject studied a subset of 12 of the 24 rule sequences during training. Subjects received 30 training blocks of 24 trials each, over a three day period. During the third day they also performed 10 transfer blocks which contained three separate trial types: old/old trials (previously solved rule sequences with previously viewed instances), old/new sequences (previously solved rule sequences with new instances), and new/new trials (new rule sequences with new instances). Subjects were forty undergraduate students at the University of Utah. During training, sequences were selected in such a way that each subject received equal amounts of practice with each individual rule in each serial position in the three sequence chain of rules. During training blocks, subjects were encouraged to go as fast as possible, while maintaining a 90% accuracy rate. During transfer, subjects were told to go as fast as possible, while maintaining 100% accuracy. Subjects were also instructed that they could retake any transfer trial they thought they might have responded to incorrectly by pressing the spacebar on the computer keyboard before responding to the following trial. The data on undetected errors is not reported here, as this part of Study III served as the pilot for developing the methodology for "detecting undetected error" (see above). Finally, subjects were shown a list of the 24 possible rule sequences (e.g., LAST-MIDPOINT-SAME) and asked to choose the 12 they had been exposed to during training.

Study III Results. Latency data from transfer established a reliable difference between old/new and new/new trials, indicative of a general sequence memory effect and composition. Latency data from transfer also revealed a non-significant trend (p=.10) toward a difference between old/old and old/new

trials, likely indicative of a weak instance effect and proceduralization. Error data mirrored the latency findings. The difference between old/new and new/new trials was significant (i.e., general sequence memory and composition), and the difference between old/old and old/new trials approached significance (p=.08; i.e., instance effect and proceduralization). Thus, in the current task environment there is strong evidence for general sequence memory and composition, and some evidence for a weaker effect of instance memory and proceduralization.

Finally, subjects' performance on identifying the rule sequences seen during training was at chance levels. This is in accord with their own anecdotal reports. Subjects seemed to have no conscious access to their memory for general rule sequences, as would be predicted by the process of composition.

Studies I, II, and III have been submitted for publication in the *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Study IV and V Objectives. The next pair of studies investigated in more detail subjects' ability to consciously access their general sequence memory and their instance memory. Study IV investigated general sequence memory, while Study V investigated instance memory. We hypothesized that subjects would have no conscious access to general sequence memory, because this memory is contained in composed productions in procedural memory. Once a production is composed, the contents of it are no longer open to conscious introspection (Anderson, 1983). Likewise, we predicted that subjects would have no conscious access to their instance memory, because it also is part of procedural memory.

Study IV Methods. Methods for Study IV were similar to those for Study III. Subjects received 45 blocks of training (each block being 24 trials long) over four sessions. Training trials consisted of 12 of the 24 possible rule sequences used in full fledged number reduction. Training rule sequences were selected so that each rule appeared in each serial position with equal frequency. Subjects also received 12 blocks of transfer trials during the fourth session. On these trials subjects were instructed to respond "old" if a number sequence represented a sequence of rule operations they had used during training, and "new" if not. A further manipulation concerned solving the items. On half the training blocks, subjects solved the items prior to making "old/new" judgments. On the other half of trials, subjects simply made "old/new" judgments without solving the items. Blocks that required solving the items were alternated with blocks that did not require solving the items. The type of block that started the transfer session was counterbalanced across subjects.

Twenty-four University of Utah students participated in Study IV.

Study IV Results. During transfer, recognition performance on the "old/new" judgment was at chance. Performance of an item was not found to affect recognition. Subjects were correct 52.50% of the time when performing items, and correct 52.57% of the time when not performing the items. Neither number significantly differed from 50%. Thus Study IV's results confirmed those of Study III: general rule sequence memory (i.e., composition) does not appear to be open to conscious access. Instead, it appears to be implicit in nature.

Study V Methods. Methods for Study V were similar to those for Study IV, except that the "new/old" judgment was based on whether the item was actually presented during training (i.e., was an old instance).

Study V Results. Study V has been run, and the data are currently being analyzed.

The first five studies reported investigated evidence consistent with Anderson's ACT* theory of skill acquisition: evidence for composition, proceduralization, and the implicit nature of these processes. The next pair of studies investigated the phenomenon undetected error making.

Study VI Objectives. Study VI contrasted high and low skill individuals on the number reduction task. It was hypothesized that high skill subjects would show greater evidence of composition and

proceduralization. Furthermore, high skill subjects should make more errors on trials that resemble previously practiced rule sequences. This would be due to a partial match between the conditions for a well-learned, strong, composed production from training (with an extremely low firing threshold), and a new sequence that is similar to the well-learned production in its initial rules (these productions would be weak, with a relatively high firing threshold).

Study VI Methods. Seventy-two University of Utah students participated in Study VI. Subjects were divided into two groups: high skill (n=38) and low skill (n=34). During training both groups received practice on 12 of the 24 possible rule sequences (balanced for frequency of occurrence of each rule in each serial position). The high skill group received 55 blocks of 24 trials each of training over five sessions. The low skill group practiced on an unrelated computer task for three sessions. They then received 15 blocks of training over two sessions.

During all the last five blocks of training (i.e., training occurring on sessions one through four) subjects were encouraged to go as fast as possible while maintaining an error rate of 10%. This was done to encourage fast, skilled performance on the task. During the last session (i.e., final five blocks of training) subjects were given a new performance goal: they were told to go as fast as possible, while being entirely accurate (i.e., error rate of 0%). Because it would be difficult to achieve this goal, they were told they could "retake" any trial on which they thought they had made an error by pressing the space bar. The space bar could be pressed, and the trial retaken, any time prior to answering the following trial. It was not possible, however, to retake earlier trials. We introduced the new performance goal, and the error retake method, to allow us to distinguish detected from undetected errors.

Transfer took place during the last 10 blocks of 24 trials of the final session. These blocks were not identified to subjects as being different from the first five blocks; however, they were comprised of three different trial types: old/old (old rule sequences using old instances; 25% of each transfer block of trials), old/new (old rule sequences using new instances; 25%), and new/new (new rule sequences using new instances; 50%). By this point, subjects were well experienced using the error detection methodology.

Finally, subjects took two questionnaires. The first questionnaire asked three questions about whether or not subjects had been using the spacebar as instructed. It was identical to the questionnaire described earlier under "Detecting Undetected Errors." The second was a listing of the 24 different rule sequences (e.g., LAST-MIDPOINT-SAME). Subjects were told that only 12 of the 24 rule sequences had been practiced during previous sessions. They were asked to circle "old" or "new" for each sequence, just as subjects in Study III had.

Study VI Results. Both groups latency data were well fit by the power law of learning. First session performance (practice blocks 1-10) did not differ significantly between the two groups. Performance for the first five blocks of the final session displayed differences in latency reflective of the differing levels of practice given the two groups: high skill subjects were approximately 800 msec faster than low skill subjects. During the 10 transfer blocks, high skill subjects were again significantly faster than low skill subjects. The difference here was approximately 500 msec (low skill subjects had gained additional skill during transfer). More importantly, the difference between high and low skill subjects during transfer was a function of trial types. With regard to the difference between old/old and old/new trials (i.e., the instance effect indexing proceduralization), high and low skill subjects both showed differences on the order of 100 msec., and did not differ significantly from one another. With regard to the difference between old sequences (old/old and old/new combined) and new sequences (new/new) (i.e., general sequence effects indexing composition), high skill subjects showed a significantly greater difference than did low skill subjects. Thus, expertise did interact with skill level in the latency data, but only for general sequence memory.

Error data were divided into two types: detected errors (on which the subject had hit the space bar to retake the trail) and undetected errors (on which the subject had failed to hit the spacebar). In general, there were somewhat more undetected errors (between 4% and 6%, depending upon trial type) than

detected errors (between 2% and 4%, depending upon trial type). Only one condition significantly deviated from this general finding: high skill subjects made more undetected errors (approximately 10%) on new/new trials. This confirmed our hypothesis that new sequences, which resembled old sequences, would lead to greater numbers of undetected errors -- but only for subjects with a high skill subjects whose productions had become composed.
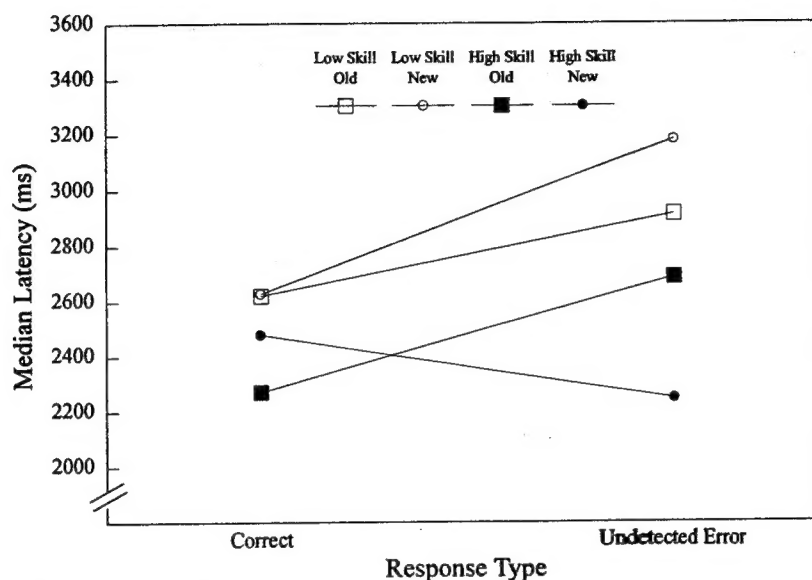
High n=25; Low n=29

Another interpretation that could explain these data, and one proposed by Anderson (1989), is that subjects may inappropriately apply weak method solutions (e.g., mistaken analogies to previous problems), which may lead to undetected errors. To distinguish our hypothesis from this explanation, we analyzed latency to errors in Study VI (see figure above). Our hypothesis predicts fast responses to undetected errors by high skill subjects who are misfiring composed, skilled productions due to a partial match of conditions. In contrast, Anderson's misapplication of weak methods predicts relatively slow responses leading to undetected errors, even among high skilled individuals. Latency data supported our prediction: high skill subjects responded as quickly on undetected errors on new/new trials as they did on correct trials. Low skill subjects, however, responded much slower on these errors than they did on correct trials, presumably because they had no composed productions to guide processing. This indicates that the nature of these error differed as a function of skill level on the task. Another interesting finding from the error latencies was that high skill subjects were relatively slow on errors made to old trials (compared to correct responses). This may reflect instances in which subjects revert from using composed, skilled memory representations to older, more error prone processing.

Data from the questionnaire asking subjects if they had used the spacebar to correct errors indicated that the vast majority of subjects had conformed to the experimental instructions. In the low skill group, 81% reported using the spacebar to correct all errors they were aware of; in the high skill group, 63% reported using the spacebar to correct all errors. The difference between groups is not significant. Of those who did not use to spacebar to correct all errors, the estimated number of uncorrected errors ranged from 2 to 10 in the low skill group, and 2 to 24 in the high skill group (again, not significantly different between the groups). If one removes those subjects who failed to correct 7 or more errors (two high skill Ss and two low skill Ss) and reanalyzes the data, the results remain the same.

Data from the rule sequence recognition test showed a slight bias toward responding "old." However, the high and low skill groups did not differ from each other in their recognition performance as measured by *d'*. Thus, as in earlier studies, general processing sequence information appears to be implicit in nature.

Study VII Objectives. Study VII was designed to test the "partial match" hypothesis in greater detail. It was argued in Study VI that a partial match between the conditions of a new rule sequence and a strong-but-wrong old sequence resulted in the inappropriate firing of the strong-but-wrong production. Latency data from high skill errors supported this interpretation; the errors were as fast as correct trials. It is possible, however, that subjects learned performance timing information independently of the rule sequence information (see, for instance, MacKay, 1982, 1987). Since all new sequence in Study VI were matched with old sequences in their first two rules, it was not possible to rule out this competing explanation.

In Study VII subjects learned only eight rule sequences during training. During transfer, they were exposed to all 24. Eight of these were old. Of the 16 new rule sequences, eight matched the old sequences in their first two rules (just as they had in Study VI). These were designated *partial-match new* sequences. The remaining eight began with an initial two rule sequence that had not been experienced during training. These were designated *mismatch new* sequences. If partial matching of conditions was the reason for the errors found in Study VI, then we should find fast errors in the partial-match new condition, but not in the mismatch new condition.

Study VII Methods. Methods, procedures, and the experimental task were similar to those in Study VI. Subjects were 49 University of Utah undergraduates. All subjects performed four sessions, with transfer blocks being presented during the last session. All subjects, therefore, were "high skill." Subjects received 10 training blocks during session one, 20 training blocks during session two, 20 training blocks during session three, and 5 training blocks during session four. During session four, the performance goal was changed from 90% accuracy to 100% accuracy, as in previous studies. Also during session four, retaking of trials by pressing the spacebar was introduced. The final 16 blocks of session four were transfer. Each transfer block consisted of the following: old sequences/old instances (33%), old sequences/new instances (33%), partial-match new sequences/new instances (17%), and mismatch new sequences/new instances (17%).

Study VII Results. Latency data showed a small, but reliable, instance effect on the order of 50 msec (difference between old/old and old/new). General sequence effects, as in previous studies, were much larger (and also reliable). When considering partial-match new trials, the sequence effect (partial match new/new versus old/new) was approximately 150 msec, with the partial match new trials being slower. When considering mismatch new trials, the sequence effect (mismatch new/new versus old/new) was approximately 400 msec, with the mismatch new trials being slower. There was a 250 msec difference between the partial-match new/new and the mismatch new/new, with the mismatch new/new being slower. Thus although previous studies showed negative transfer with regard to errors in the partial match new situation, there is also positive transfer with regard to latency, presumably due to the overlap of the first two rules.
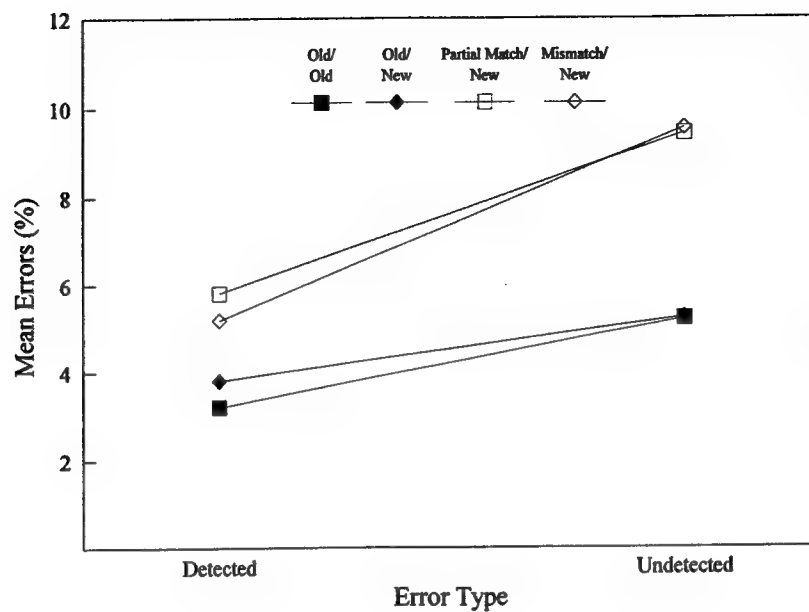
Error data are presented for the four conditions above. As can be seen from the figure, subjects made slightly more detected errors on partial-match new and mismatch new, as compared to old sequences. The partial-match and mismatch conditions did not differ from each other, however. Undetected errors showed the same pattern, but the effects were even stronger.

N=34
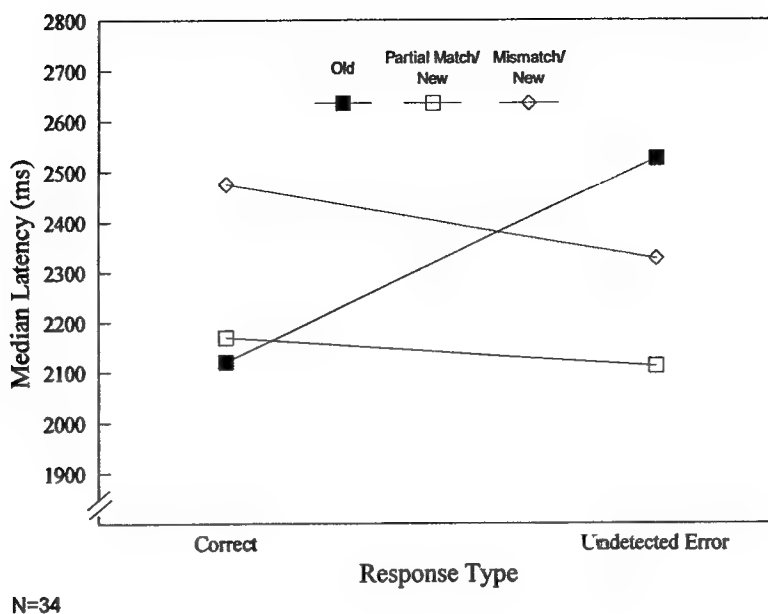
The latency data for errors, which is the primary analysis of interest in Study VII, is displayed above. The two old conditions (old/old and old/new) were combined in this analysis. The analysis also includes only those individuals who made at least two undetected errors in all trial conditions. First, note that the latency for partial match new undetected errors is fast, just as in Study VI. It is as fast as correct old trials and correct partial mismatch trials. This is consistent with the interpretation made in Study VI: undetected errors in the partial match condition appear to be due to the misfiring of strong-but-wrong composed productions that partially overlap with the rule sequences required for correct solution of these items.

Second, note that the latency for undetected errors on old sequences is quite slow. This finding was unexpected, but replicates a finding in Study VI. These errors seem consistent with what Reason (1990) has described as *overattention*. Such errors result from a performer intervening and consciously controlling performance, when, in fact, he or she would have been better off to allow skilled memory representations to guide performance. An example would be a baseball batter in a slump, who finds that his batting deteriorates even further as he tries to consciously "correct" his swing.

Studies VI and VII have been submitted for publication in the *Journal of Experimental Psychology: General*.

Future Research. Our studies have been consistent in suggesting that undetected errors can result from long-term priming due to the inappropriate firing of strong-but-wrong composed productions. We plan to extend this research by attempting to produce undetected errors due long-term priming in a different task domain. We are currently programming a new task, denoted *procedural learning*, in which subjects learn to classify numbers according to a set of rules which can be represented as a decision tree. Subjects must decide whether a number should appear on the top or the bottom of the computer screen, and compare that to the actual position in which the number is presented. They respond either "L", for "like," if the number is presented in its correct screen location, or "D", for "different," if the number is presented in the incorrect screen location. The dimensions on which number are evaluated are: (a) presentation mode (either as an English word or an Arabic numeral); (b) number magnitude (either large or small, defined arbitrarily and

consistently within an experiment); (c) divisibility by two (numbers may be either odd or even); and (d) presentation position (either top or bottom of the computer screen). At first, this task is performed in slow controlled fashion by subjects, but they quickly proceduralize their knowledge and perform in a fast and effortless fashion. Procedural learning experiments will take place during year two of the grant period.

<div align="center">Short-Term Priming</div>

The first seven studies established undetected errors in a training/transfer paradigm where training was restricted to a subset of the total universe of possible rule sequences in the number reduction task. Short-term priming studies were aimed at investigating the possibility that undetected errors due to strong-but-wrong composed productions could be induced over a small number of trials within a single block. The rationale was similar to that for long-term priming: if either the repeated firing of a production temporarily strengthens it, or if the presentation of contextual information associated with the firing of a production spreads activation that temporarily lowers the firing threshold of the production, then temporary priming should be possible. Predictions were similar to those for long-term priming: we would expect errors in situations where a temporarily primed production shared a partial match to the conditions of an item.

Study VIII Objectives. Study VIII investigated whether temporary priming could be found in the complex version of the number reduction task. Subjects became highly practiced on the all 24 of the possible three rule sequences (but different instances). During transfer, they encountered groups of one, two, and three consecutive trials using the same rule sequence, and then were switched to a partial match sequence that differed only in the final rule. The question of interest is whether such "short-term primed" targets would show larger numbers of undetected errors than matched trials that were unprimed.

Study VIII Methods. Thirty-two University of Utah students served as subjects in Study VIII. Three were eliminated because their data indicated lack of effort, leaving 29 students in the final data analysis. Subjects received five training sessions of 12 blocks each, with each block containing 24 trials. This allowed each of the 24 possible rule sequences to appear once in each block. Each presentation of a rule sequence used a unique instance of that sequence. The final (sixth) session was a transfer session containing 24 blocks of 30 trials each. The first six trials in each transfer block served as "warm-up" trials, and contained no priming pattern. The final 24 trials were divided into six groups of four trials each. Half of these were "short-term primed" groups, and half were "short-term unprimed" controls. In a "short-term primed" group, whichever rule sequence was assigned as a target (e.g., SAME-MIDPOINT-LAST) would be preceded by either one, two, or three partial match rule sequences (e.g., SAME-MIDPOINT-CONTIGUOUS). Each transfer block contained one one-prime, one two-prime, and one three-prime grouping of four trials. Thus three rule sequences received short-term priming in each transfer block. Three other rule sequences were assigned as "short-term unprimed" targets for the other three groups of four trials. These rule sequences were not preceded by partial match priming trials. We tested our hypotheses about short-term priming by comparing target trial errors following primed versus unprimed sequences. Primed and unprimed groups of four trials alternated during transfer blocks. As in previous studies, subjects could use the spacebar to correct errors during transfer.

Study VIII Results. Latency data revealed no significant differences between short-term primed and unprimed trials. Likewise, number of priming trials preceding a target was not related to latency. Similar results were found for number of undetected errors: priming and number of priming trials were not related to number of undetected errors.

Study IX Objectives. In Study VIII, subjects received practice on all 24 possible rule sequences during each training block. It is possible that subjects composed not only the three rules in each sequence during training, but also the process of switching from one rule sequence to another (e.g., see Carlson and Yaure, 1990). This would have minimized the likelihood of undetected errors during transfer, as subjects had become skilled at switching from one rule sequence to another. To minimize this learning, Study IX replicated Study VIII, except that training was blocked by rule sequence. Thus, in any given training block, subjects practiced only one rule sequence.

Study IX Methods. Methods for Study IX were similar to those for Study VIII, except for the following changes: (a) training was blocked by rule sequence, so that only a single rule sequence was practiced per training block; (b) the number of training sessions was reduced from five to four, and the number of transfer blocks was increased from one to two (this was done to increase the number of opportunities for making undetected errors); and (c) the last rule, which stated that if two digits differed by more than two, they could be reduced to the last of the two digits, was replaced by the first rule, which stated that if two digits differed by more than two, they could be reduced to the first of the two digits. This last change was required by the blocking procedure. When the last rule occurred in the final position (e.g., SAME-CONTIGUOUS-LAST), problems could be solved by simply reporting the final digit of the problem. With the first rule, subjects could not use this strategy, since the first digit of the final two could be modified by intermediate solutions to the first two rule applications within a problem.

Nineteen University of Utah students served as subjects in Study IX. None were eliminated due to lack of effort; thus, 19 subjects were included in the final data analysis.

Study IX Results. Once again, both latency and error data showed no effect of short-term priming.

Study X Objectives. Studies VIII and IX found no evidence of undetected errors due to short-term priming. The rationale for these experiments was straight forward, and the priming of target rule sequences was direct. In Study X, we explored short-term priming of an indirect nature. We reasoned that it might be possible to lower the firing threshold for a composed production by presenting contextual cues present during the composition of that production. If this contextual information was later (i.e., during transfer) presented inappropriately with a partially matching rule sequence, it could lead to an undetected error due to firing of the strong-but-wrong production associated with the contextual cue.

Study X Methods. The task in Study X was the complex number reduction task. Subjects practiced eight rule sequences during training: four randomly chosen sequences and four corresponding partial match sequences (matching their mates in the first two rules). The experiment took place over four experimental sessions. During session one, subjects received 12 training blocks; during session two, 18 training blocks; during session three, 18 training blocks; and during session four, 3 training blocks. The remaining 15 blocks of session four were transfer blocks. Each block in training and transfer was 32 trials long. As in previous studies, error detection was introduced during the final session.

A major difference between Study X and previous studies concerned the presentation of the digit strings to be reduced. In previous studies these had been presented in the center of the computer screen against a black background. In Study X, each of the eight rule sequences was presented in a different spatial location (at either 0, 45, 90, 135, 180, 225, 270, 315 degrees of orientation from vertical, approximately 2 inches from the center of the screen) and in a different color. Sequences were oriented so that their partial match mates were opposite them in terms of presentation position (i.e., sequence A at 0 degrees, and partial match A at 180 degrees) and presented in the complementary color. Presentation position and color cues were perfectly correlated with rule sequences during training. During transfer, two types of trials were possible: (a) "switched" trials, in which the presentation position and color of a rule sequence were switched to its partial match mate (25% of trials), and (b) consistent trials, which maintained the mapping of presentation position and color to rule sequence learned during training (75% of trials). Each rule of the eight rule sequences was switched once during each transfer block. We hypothesized that undetected error would be more numerous on switch trials, due to short-term contextual priming.

Study X Results. Error data did not support an increase in number of undetected errors on "switch" trials. Thus, there was no evidence in favor of short-term priming due to indirect or contextual cueing.

Future Research. The failure to find evidence of undetected errors in short-term priming situations was surprising, both in light of the success of producing undetected errors through long-term priming and anecdotal reports of errors due to such short-term priming phenomena. We are currently planning two

future studies on short-term priming. The first involves trying a different experimental task (procedural learning), and the second involves investigating short-term priming in low skill (i.e., relatively novice) subjects.
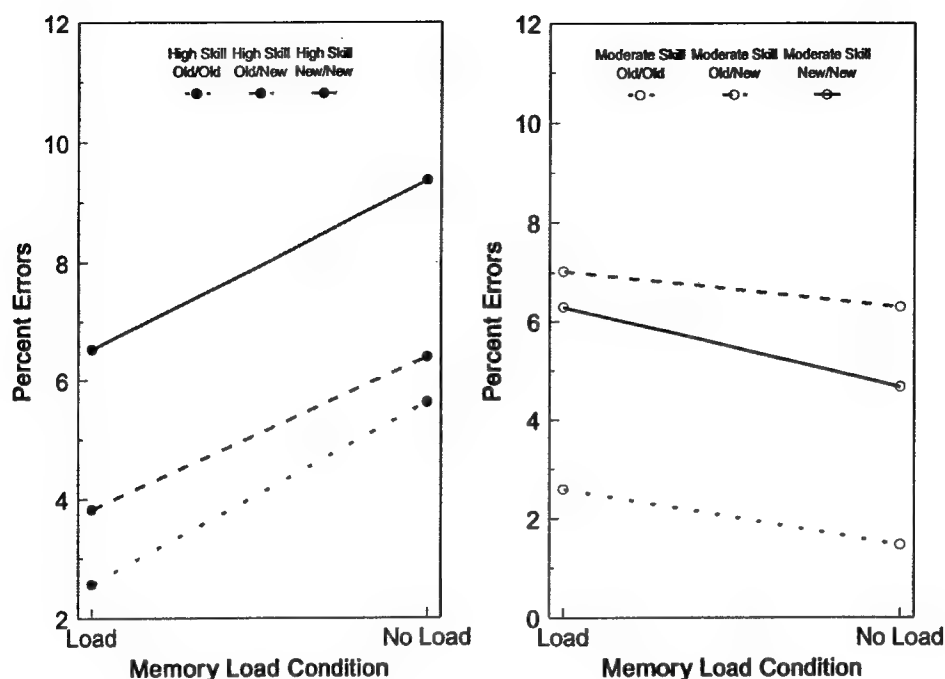
Working Memory Load

A consistent theme in the writings of theorists concerned with error making is that a working memory load leads to errors (e.g., Heckhausen & Beckman, 1990; Norman, 1981; Reason, 1977, 1990). In general, the theories say the following: if an individual is engaged in a task that they have routinized (i.e., are highly skilled at), and she or he is also engaged in second activity that is not routinized (so that it consumes working memory), this individual is susceptible to the intrusion of strong-but-wrong processes that are competing with the routinized activity. An example would be making tea while holding a conversation. One might inadvertently put instant coffee in the cup rather than a tea bag, resulting in a cup of coffee rather than tea. The strong-but-wrong coffee process intruded on the weaker tea processes, due, in part, to the load put on working memory by holding a conversation simultaneously with making a hot beverage.

Anderson's (1987) ACT* theory, however, makes a different set of predictions. Anderson predicts that a working memory load will be more detrimental to a moderately skilled individual than a highly skilled one. This is because the highly skilled individual has composed her productions, thus freeing up working memory capacity to deal with the secondary task. The moderately skilled individual requires more productions to deal with the primary task, and therefore has fewer attentional resources to deal with the secondary task. Once attentional resources become overloaded, either the primary task or the secondary task, or both, will become error prone. To extent that both moderately and highly skilled individuals are using procedural memory to guide their performance, errors may be undetected.

Study XI Objectives. A pilot study was devised to investigate to effect of a working memory load on undetected errors due to the intrusion of strong-but-wrong productions. Two groups of subjects were practiced on the complex version of the number reduction task. One group of subjects a moderate amount of practice on the task, while the other group was highly practiced. Transfer blocks alternated between those requiring subjects to engage in a secondary task (i.e., memory load condition) and those without a secondary task (i.e., no memory load condition). We hypothesized that subjects would make greater numbers of undetected errors with a memory load, and that moderately skilled subjects would show a larger effect for memory load.

Study XI Methods. University of Utah students served as subjects for Study XI. Nine subjects were randomly assigned to the moderate skill group, and 13 subjects were randomly assigned to the high skill group. Session one was identical for both skill groups: 10 blocks of training on the number reduction task. During sessions two, three, and four the high skill group received an additional 15 blocks of number reduction practice per session, while the moderate skill group participated in a different experimental task that involved computerized presentation of stimuli and data collection. During session five, both groups received five additional training blocks during which error correction was introduced and the accuracy criterion changed from 90% to 100%. Both groups then received 10 blocks of transfer containing three types of trials: old rule sequences/old instances (25% of trials); old rule sequences/new instances (25% of trials); and new partial match sequences/new instances (50% of trials). For half of the 10 transfer blocks, subjects were also required to perform a secondary task designed to produce a working memory load. This task consisted of keeping track of how many times to randomly selected digits had served as answers to the number reduction problems within that block. Working memory load blocks alternated with blocks that did not require the secondary task (i.e., no working memory load blocks).

15

Figure: Two panels plotting Percent Errors against Memory Load Condition (Load, No Load). Left panel — High Skill conditions (Old/Old, Old/New, New/New). Right panel — Moderate Skill conditions (Old/Old, Old/New, New/New).

Study XI Results. Results, in terms of percentage of undetected errors for each trial type, are presented in the figure above. The left panel presents data from the high skill condition, while the right panel presents data from the moderate skill condition. First, it is clear that memory load has an effect on undetected error making. However, the direction of this effect is a function of skill level. Highly skilled subjects actually benefit from a memory load, while moderately skilled subjects suffer a cost due to the load. While we predicted the detriment to performance in the moderately skilled condition, the results from the highly skilled conditions were surprising. We hypothesize that the memory load, which involved numbers and was highly distracting, forced highly skilled subjects to abandon their composed productions and rely on more controlled processing for task solution. This eliminated the intrusion of partially matching, strong-but-wrong productions, which we hypothesized to be a major factor in producing undetected errors. Thus, performance for this group improved in the memory load condition. Moderately skilled individuals, however, found their working memory capacity exceeded in the memory load condition, and began making errors as Anderson (1987) predicted.

Future Research. Future research needs to replicate the findings of Study XI in a larger sample, and to extend these findings to other task domains. Our hypotheses concerning the mechanisms for the divergence of error making under working memory load conditions need to be tested in greater detail as well.

General Directions for Future Research

Temporary Changes in Procedures. One situation which may produce undetected errors is the temporary change in a well learned procedure. For instance, if one normally edits on a particular word processor, but must make changes to a document on a similar, yet slightly different, word processor, several well learned routines for editing may have to be modified. The modifications are only temporary, because you intend on returning to your word processor of choice as soon as possible. Can composed productions be modified

easily, or temporarily disregarded in favor of controlled processing, without undetected errors arising from strong-but-wrong intrusions?

We plan to study temporary changes in procedures using a task pioneered by Elio (1986; Frensch, 1991) that involves computing a water quality index by following a set of computational steps. These steps involve routine arithmetic operations on a set of data presented to the subject. During transfer, subjects will be required to change one of the component steps, but only during certain blocks of trials. The frequency of undetected and detected errors will be measured, as well as latency to make these errors. We hypothesize undetected errors to occur as a function of either strong-but-wrong intrusions or due to the loss of a subsidiary goal (i.e., to change the procedure). Latencies should help decide which process is causing the errors.

The water quality task is currently being programmed.

Individual Differences in Susceptibility to Undetected Error Making. From both a theoretical and personnel/management perspective, knowing who is particularly susceptible to undetected error making is of interest. Tests, such as the Cognitive Failures Questionnaire (Broadbent, Cooper, Fitzgerald, & Parkes, 1982) and Reason & Mycielska's (1982) Absentmindedness questionnaire, exist which purport to measure an individual's likelihood of making slips. However, we have collected pilot data which indicate the following:

(1) These questionnaires, while correlated with each other, are unrelated to undetected error making in our experimental tasks.
(2) Other personality measures, such as impulsivity, need-to-achieve, cognitive structure, are also unrelated to both undetected error making, and the questionnaires that purport to measure slips.
(3) Working memory ability is related to undetected error making, and unrelated to questionnaires that purport to measure slips.

These finding need to be confirmed with larger samples, but they are potentially very important. Paper and pencil tests which claim to measure error making appear only measure people's opinions about their errorfulness. These opinions appear to be reliable (if not valid). Other personality measures, which are defined more broadly, do no better. But working memory, which has been traditionally associated with ability measures (Kyllonen & Christal, 1990), rather than personality measures, is a successful predictor. Presumably size of working memory limits the effectiveness of procedural skill and the ability to monitor errors.

Computer Simulation of Partial Match Error Making. Occasionally theories of skill acquisition make predictions that, on the face of them, seem quite similar (e.g., Anderson, 1983; MacKay, 1987). However, when the theories are completely specified, as in a computer simulation, their predictions diverge. We have long believed that simulating undetected error making would allow us to better understand which theoretical position is more correct, as well as to better understand our own data. We would, in the future, like to simulate the partial match mechanism which we have postulated is the basis for many undetected errors. This has not been successfully accomplished by other researchers, to the best of our knowledge (e.g., Anderson, 1993).

Conclusion. During our first year, we have learned a great deal about undetected error making and its underlying mechanisms. During our second year we hope to generalize these results to new task domains, and to explore new questions of interest.

References

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review, 94*, 192-210.

Anderson, J. R. (1989). The analogical origins of errors in problem solving. In D. Klahr & K Kotovsky (Eds.), *Complex information processing* (pp. 343-372). Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Broadbent, D. E., Cooper, P.F., Fitzgerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology, 21*, 1-16.

Bryan, W. L., & Harter, N. (1899). Studies on the telegraphic language: The acquisition of a hierarchy of habits. *Psychological Review, 6*, 345-375.

Carlson, R. A., & Yaure, R. G. (1990). Practice schedules and the use of component skills in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 484-496.

Crossman, E. R. F. W. (1959). A theory of the acquisition of speed-skill. *Ergonomics, 2*, 153-156.

Elio, R. (1986). Representation of similar well-learned cognitive procedures. *Cognitive Science, 10*, 41-73.

Fitts, P. M., & Posner, M. I. (1967). *Human Performance*. Belmont, CA: Brooks Cole.

Frensch, P. A. (1991). Transfer of composed knowledge in a multi-step serial task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 997-1016.

Heckhausen, H., & Beckman, J. (1990). Intentional action and action slips. *Psychological Review, 97*, 36-48.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14*, 389-433.

LaBerge, D. (1973). Attention and the measurement of perceptual learning. *Memory & Cognition, 1*, 268-276.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492-527.

MacKay, D. G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review, 89*, 483-506.

MacKay, D. G. (1987). *The organization of perception and action*. New York: Springer-Verlag.

Norman, D. A. (1981). Categorization of action slips. *Psychological Review, 88*, 1-15.

Reason, J. T. (1977). Skill and error in everyday life. In M. Howe (Ed.), *Adult learning: Psychological research and applications*. London: Wiley.

Reason, J. T., & Mycielska, K. (1982). *Absentmindedness? The psychology of mental lapses and everyday errors*. Englewood Cliffs, NJ: Prentice-Hall.

Reason, J. T. (1990). *Human error*. Cambridge, UK: Cambridge University Press.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing. I. Detection, search, and attention. *Psychological Review, 84,* 1-66.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing. II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127-190.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Appendix A
Publications and Submissions

Woltz, D. J., Bell, B. G., Kyllonen, P.C., & Gardner, M. K. (under review). *Memory for order of operations in the acquisition and transfer of sequential cognitive skill.* Manuscript submitted to *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Woltz, D. J., & Christal, R. E. (under review). *Detected and undetected errors in working memory.*

Ackerman, P.A., & Woltz, D. J. (in press). Determinants of learning and performance in an associative memory/substitution task: Constraints, individual differences, and volition. *Journal of Educational Psychology.*

Woltz, D. J., & Shute, V. J. (in press). The time course of forgetting exhibited in repetition priming of semantic comparisons. *American Journal of Psychology.*

Gardner, M. K., & Sternberg, R. J. (1994). Novelty and intelligence. In R. J. Sternberg & R. K. Wagner (Eds.), M*ind in context: Interactionist perspectives on human intelligence.* Cambridge, UK: Cambridge University Press.

Gawlick-Grendall, L. A., & Woltz, D. J. (1994). Dominance of meanings for 120 homographs. *Behavior Research Methods, Instruments, & Computers, 26,* 5-25.

Woltz, D. J., & Shute, V. J. (1993). Individual difference in repetition priming and its relationship to declarative knowledge acquisition. *Intelligence, 17,* 333-359.

Appendix B
Participating Professionals

Dan J. Woltz, Ph.D., Assistant Professor, Co-Principal Investigator
Michael K. Gardner, Ph.D., Associate Professor, Co-Principal Investigator

Janet G. Madsen, Graduate Student (Ph.D.)
      University of Utah Graduate Research Fellowship, 1993-1994
Brian G. Bell, Graduate Student (Ph.D.)
      WPA Student Research Award, 1993
James Farnham, Graduate Student (Ph.D.)
      University of Utah Graduate Research Fellowship, 1993-1994
Susan Curtis, Graduate Student (M.S.)

Appendix C
Papers and Presentations

Madsen, J. G., & Woltz, D. J. (April, 1995). *Changes in access to verbal knowledge as procedural skill develops*. Paper accepted for presentation at the annual meeting of the American Educational Research Association, San Francisco, CA.

Woltz, D. J., Gardner, M. K., Bell, B. G., & Farnham, J. M. (July, 1994). *Memory for processing sequences can produce undetected errors in skilled performance*. Paper accepted for presentation at the annual meeting of the American Psychological Society, Washington, D.C.

Woltz, D. J., Gardner, M. K., & Bell, B. G. (April, 1994). *Knowledge compilation and undetected errors by skilled performers*. Paper accepted for presentation at the annual meeting of the American Educational Research Association, New Orleans, LA.

Gardner, M. K., Woltz, D. J., & Bell, B. G. (August, 1993). *The costs of expertise: Undetected errors in cognitive skill acquisition*. Paper presented to US Air Force Armstrong Laboratories, Brooks AFB, San Antonio, TX.

Woltz, D. J. (June, 1993). *Individual differences in implicit and explicit memory for procedural sequences in skill acquisition*. Paper presented at the annual meeting of the American Psychological Society, Chicago, IL.

Bell, B. G., Woltz, D. J., & Madsen, J. G. (April, 1993). *The role of memory for procedural sequences in skill acquisition*. Paper presented at the annual meeting of the Western Psychological Association, Phoenix, AZ.

Appendix D
New Discoveries, Inventions, and Patents

None